



**UNIVERSIDADE ESTADUAL DE CAMPINAS
SISTEMA DE BIBLIOTECAS DA UNICAMP
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELECTUAL DA UNICAMP**

Versão do arquivo anexado / Version of attached file:

Versão do Editor / Published Version

Mais informações no site da editora / Further information on publisher's website:

<https://genus.springeropen.com/articles/10.1186/s41118-020-00105-3>

DOI: 10.1186/s41118-020-00105-3

Direitos autorais / Publisher's copyright statement:

©2020 by SpringerOpen. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

ORIGINAL ARTICLE

Open Access



Temporal and spatial trends of adult mortality in small areas of Brazil, 1980–2010

Bernardo L. Queiroz^{1*} , Everton E. C. Lima² , Flávio H. M. A. Freire³ and Marcos R. Gonzaga³

* Correspondence: lanza@cedepplar.ufmg.br

¹Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627–Cedeplar/Face/UFMG, Belo Horizonte 31270-901, Brazil
Full list of author information is available at the end of the article

Abstract

To determine the variations and spatial patterns of adult mortality across regions, over time, and by sex for 137 small areas in Brazil, we first apply TOPALS to estimate and smooth mortality rates and then use death distribution methods to evaluate the quality of the mortality data. Lastly, we employ spatial autocorrelation statistics and cluster analysis to identify the adult mortality trends and variations in these areas between 1980 and 2010. We find not only that regions in Brazil's South and Southeast already had complete death registration systems prior to the study period, but that the completeness of death count coverage improved over time across the entire nation—most especially in lesser developed regions—probably because of public investment in health data collection. By also comparing adult mortality by sex and by region, we document a mortality sex differential in favor of women that remains high over the entire study period, most probably as a result of increased morbidity from external causes, especially among males. This increase also explains the concentration of high male mortality levels in some areas.

Keywords: Small areas, Mortality estimation, Demographic techniques, Spatial analysis

Introduction

Over the last several decades, Brazil has experienced an accelerated decline in infant, child, and adult mortality, leading to a roughly 15-year median gain in life expectancy at birth from 1950 to 2010, a much greater increase than in developed countries over the same period. Nonetheless, despite signs of convergence in infant mortality because of fewer deaths from infectious diseases, large differences remain across Brazilian regions in both life expectancy at birth and adult mortality (França et al. 2017). Yet although many studies describe such variations in infant and child mortality for Brazil's small areas (Souza et al., 2010; Barufi et al., 2012; McKinnon, 2010), by providing no substantive evidence on spatiotemporal trends in adult mortality, they ignore an issue that complicates most such research in less developed countries; namely, the overall quality of mortality data across regions (Hill, 2017; Murray et al. 2010; Hill et al. 2009).

In Brazil, this data quality problem is characterized by incomplete coverage of vital registration systems, errors in age declaration for both population and death counts, and a lack of reliable information on cause of death (Queiroz et al. 2017; Borges, 2017; Gonzaga and Schmertmann, 2016; Lima and Queiroz, 2014; França et al. 2008; Paes, 2005; 2007; Luy, 2011). Given that successful health planning requires disaggregated mortality measures that accurately reflect regional health and morbidity variations, this lack of reliable estimates also has a negative impact on public policy (Fenelon, 2013; Ram et al. 2015; Ruther et al., 2017; Divino et al., 2009; Tsimbos et al., 2014). Proper mortality estimates at the local level are also essential for the identification of more vulnerable populations, the development of strong public health policies across regions, and the achievement of Sustainable Development Goals (SDGs).

In this paper, therefore, we address this research gap by first applying TOPALS to Ministry of Health mortality data for 137 small Brazilian areas to estimate and smooth adult mortality (45q15) rates across age groups and then employing death distribution methods (Hill et al., 2009) to assess the quality of the death registration data. Based on these estimates, we are then able to identify the regional variations and spatial patterns of adult mortality over time and by sex between 1980 and 2010. We find that, overall, mortality data has improved over the years in Brazil, with mortality coverage increasing from roughly 80% in 1980 to over 95% in 2010. This improvement has also spread to Brazilian regions, resulting in a large portion of the country demonstrating nearly complete death count coverage. Brazil has also experienced changes in data quality, with a decline in deaths registered as “ill-defined” (less than 6% in 2017) and a decrease in missing age and sex information on death certificates from 1% in 1980 to 0.3% in 2010 (albeit with large regional variations). All these improvements can be attributed primarily to public investment in the collection of health data. At the same time, however, because of an increased number of deaths from external causes, particularly among males, mortality differences by sex have remained substantial, accounting for the high concentration of male adult mortality in some areas of the country.

Mortality studies in small areas

To illustrate the importance of regional mortality and health knowledge for informing public health policies and health system coverage, Kulkarni et al. (2011) demonstrate the large variation in life expectancy between US regions, with some presenting much higher mortality rates than those observed in other developed countries. Similar research by Kibele et al. (2015) for Germany further reveals that the factors that explained such variation in the past continue to influence current regional variation. Hence, as in other developed regions, small areas in Brazil are likely to be characterized by huge mortality differentials at a time when changing demographics are increasing the demand for appropriate local health policies (França et al. 2017).

At the same time, demographers and health experts’ increasingly greater access to geocoded data has augmented the importance of accurately estimating vital rate schedules for small areas despite the inherent challenges. These latter include not only the problem of country-level data limitations from small population size and event numbers (Alexander et al. 2017; Gonzaga and Schmertmann, 2016; Lima et al. 2016; Assunção et al. 2005), but the unstable rate estimates common for small areas with small risk populations even in national census or other very large datasets. As a result, traditional

demographic techniques often yield extreme estimates for small population areas dominated by sampling noise that may have little relation to underlying local mortality risks (Alexander et al. 2017; Gonzaga and Schmertmann, 2016; Lima et al. 2016; Assunção et al. 2005).

To combat these challenges, an emerging body of literature on small area mortality estimation is proposing several methodological alternatives for handling vital rate estimates in less populated regions (Bernardinelli and Montomoli, 1992; Leone 2014; Ahmed and Hill 2011; Lima et al. 2016; Stephens et al., 2013; Schmertmann and Gonzaga 2018), among them, the use of empirical Bayesian models (Bernardinelli and Montomoli, 1992). Another recent approach, albeit one not focused on small areas, applies locally weighted regression functions and death distribution methods (DDM) to census data from Lesotho and Nicaragua to produce smoothed mortality estimates (Leone 2014). For less populated areas, Ahmed and Hill (2011) employ population characteristics and contextual factors to estimate maternal mortality rates in small areas of Bangladesh; in particular, fitting a random effects Poisson regression model that represents maternal death counts as their proximate determinants. Alternatively, Adair and Lopez (2018) propose a symptomatic variable method for estimating death count registration completeness, which, although easily applicable to small areas, offers no capability for age profile adjustment.

In other studies for Brazil, Gonzaga and Schmertmann (2016) adapt De Beer's (2012) TOPALS (tool for projecting age-specific rates using linear splines), which combines a relational model with Poisson regressions. Although this model was originally used to smooth and forecast mortality in Human Mortality Database countries, by extending it to Brazilian microregions, these authors are able to produce smooth mortality schedules for small areas. Similarly, Lima et al. (2016) combine a hierarchical Poisson model with DDM to smooth mortality schedules, enabling them to estimate life expectancies and ${}_{45}q_{15}$ for 862 municipalities in two Brazilian regions. In a more recent study, Schmertmann and Gonzaga (2018) draw on previous DDM estimates and a verbal autopsy by Brazilian public health experts to develop a Bayesian regression model for small-area mortality schedules that combines a relational model with probabilistic prior information on death registration coverage. This model simultaneously controls for the problems of small local samples and underreporting of deaths.

To address the specific problem of registration, Oliveira et al. (2017) develop a random-censoring Poisson model (RCPM) that accounts for any uncertainty about either the count or data reporting process. As an alternative, Alexander et al. (2017) propose a hierarchical Bayesian model that produces subnational mortality estimates based on temporal and spatial information, which when tested using simulated US county and real French *départements* data, yields reasonable results but some degree of uncertainty in the estimates (Alexander et al. 2017). Another approach to these problems is to treat each small area as a separate unit of analysis and not incorporate neighboring information or regression models based on micro-data for larger areas (Ruther et al., 2017). For example, Tsimbos et al. (2014) use a combination of model-life tables and regression models to estimate life expectancy at the local level. Yet, none of the above small area studies and methods directly addresses the issue of incomplete death count records, a failing we address in the next section by proposing an easily replicable method that could be applied to a variety of countries with defective information.

Data and methods

Brazilian vital registration and population data

Our primary data source, the Ministry of Health DATASUS (<http://www2.datasus.gov.br>), provides municipality-level information on total number of deaths and causes of death by age and sex, with mortality data organized by ICD revision code (9th revision for 1980–1995, 10th for 1996 on). Although data are available from 1979 onward, we restrict our dataset to 1980–2010, a period during which missing sex and age information reduced sharply from an average 1% per death to only 0.3%, albeit often with a considerable annual range (e.g., from 3% in Mato Grosso to under 0.5% in São Paulo). We standardize our estimation procedure for all years by applying a proportional distribution of unknown age and sex data based on the proportional share of known sex and age information in the data. As expected, we note a higher percentage of death records with unknown sex and age in smaller areas of lesser developed regions, but the quality of this information has improved steadily over time.

Geographically, Brazil contains five major regions (North, Northeast, Southeast, South, and Midwest) comprising 27 federal states subdivided into 137 mesoregions, which in turn are subdivided into 558 microregions that comprise 5565 municipalities (according to the 2010 census). These municipalities have the responsibility for death data cleaning and compilation and then transferring an electronic data file to the national office every 3 months. The Brazilian Census (1980, 1991, 2000, and 2010) is responsible for gathering population age and sex data at the local level, but although considered of good quality, an underenumeration problem may affect coverage for small areas. That is, according to post-census enumeration surveys, a 7.3% underenumeration occurred in 1980, around 4% in 1991 and 2000, and as high as 9% in 2010, with the northern and northeastern states experiencing the highest levels over time. We observed an important decline in missing information on age and sex. For the years 1980 and 1991, we have used pro-rata distribution to allocate this missing data by age and sex for all localities. The percentage of missing information has declined substantially in the next years of 2000 and 2010. Moreover, IBGE reported that the quality of the census, measured by the estimated number of underreporting, declined from 1991 onwards lower than previous years (IBGE, 2008), and they also indicate a large regional difference across regions. This issue is more relevant for smaller areas. For ease of comparison, we use the IBGE mesoregion definition to aggregate municipalities in comparable small areas, yielding a 1980–2010 data sample for 137 small areas of Brazil (Ehrl, 2017). Two major advantages of this choice are that (1) these areas did not change their boundaries over the period of analysis, and (2) they are similar in regional and socioeconomic characteristics.

Methodology

We focus our analysis on 45q15, a synthetic measure of adult mortality, with 15 as the assumed age of entry into adulthood being the turning point at which declining childhood mortality risks are replaced by increased mortality risks for young-adult age groups. Because this simple measure covers a substantive age range up to age 60, it avoids problems inherent in mortality estimates at more advanced ages while offering easy comparability with existing studies. Not only will the results help fill the research void on variation in adult mortality across Brazil's different regions, but adequate estimates of adult mortality could be used in other models to produce life table estimates for small areas.

Estimating adult mortality rates in less populated areas

The TOPALS relational model (Gonzaga and Schmertmann, 2016) used to estimate mortality rates by sex and age with a log mortality schedule that is the sum of two functions: (1) a constant schedule (standard) that incorporates basic age patterns, and (2) a parametric piecewise-linear function comprised of straight-line segments between designated ages (knots). This latter represents differences between the standard and log mortality schedule in the population of interest, as expressed in the following equation:

$$\underbrace{\lambda(\alpha)}_{100 \times 1} = \underbrace{\lambda^*}_{100 \times 7} + \underbrace{B}_{100 \times 7} \underbrace{\alpha}_{7 \times 1} \quad (1)$$

where λ is a vector of log mortality rates in a mesoregion, λ^* is a standard schedule (the national log mortality rate), B is a matrix of constants in which each column is a linear B-splines basis function with knots defined at exact ages ($x = 0, 1, 10, 20, 40, 70, 100$; cf. Gonzaga and Schmertmann, 2016), and α is a vector of parameters representing offsets to the standard schedule. In Eq. (1), the α values represent additive offsets ($\lambda_x - \lambda_x^*$) to the log mortality rate schedule at knots between which the offsets change linearly with age. According to Gonzaga and Schmertmann (2016), based on any set of observed age-specific deaths (D_x) and populations (N_x), deaths are distributed as independent Poisson variables with a log likelihood of

$$\log L(\alpha) = \text{constant} + \sum_x [D_x \lambda_x(\alpha) - N_x \exp(\lambda_x(\alpha))] \quad (2)$$

Given TOPALS' inherent properties, the choice of a standard schedule is far less important in our model than in other demographic relational models like indirect standardization, especially as the assumption that all mesoregions within any Brazilian state have the same age-specific mortality rates schedule is unreliable (Gonzaga and Schmertmann, 2016). Moreover, because TOPALS estimates parameters by maximizing a Poisson likelihood function for age-specific deaths conditional on age-specific exposure, these estimates have a 95% interval confidence, a major improvement for small-area estimation in less developed countries. As a result, rather than the point estimates of mortality measures offer by earlier studies, we can report the error term, which may be closely related to data quality, population size, and other local characteristics.

Estimating completeness of death registration for mesoregions

To obtain the expected number of deaths per population, we employ a DDM approach that evaluates death count coverage in relation to population count (Hill et al. 2009). Despite some limitations, DDM, whose formalization is extensively detailed in the literature,¹ provides highly robust and consistent results for a series of applications across the globe (Peralta et al. 2019; Hill, 2017; Wang et al. 2016). In brief, DDM derives the mortality age pattern for a set period by comparing the death age distribution with the population age distribution via three primary methods: general growth balance (GGB, Hill 1987), synthetic extinct generations (SEG; Bennett and Horiuchi, 1981), and adjusted synthetic extinct generations (SEG-adj; Hill et al. 2009). In doing so, it makes several strong

¹See, for example, Bennett and Horiuchi (1981); Hill (1987); Hill et al. (2005); Hill et al. (2009); Dorrington et al. (2014a; 2014b); Murray et al. (2010).

assumptions: the population is closed, the degree of death coverage and degree of population count coverage are constant by age, and declaration of the ages of the living and dead is free of error.

The first DDM method, GGB, is derived from the basic demographic balancing equation, whose expressed identity is that population growth rate is equal to the difference between entry and exit rates. Because this identity holds for open-ended age segments $x+$, and in closed populations, the only entries are birthdays at age x , the “birthday” rate $x+$ minus the growth rate $x+$ provides a residual estimate of the death rate $x+$. If this residual estimate can be calculated from the population data in two national censuses, then a direct comparison of recorded deaths enables estimation of death recording completeness relative to the population (Hill, 1987; Hill et al., 2005; Hill et al. 2009). That is, the relations between input rates and the difference between growth and mortality rates in every age group allow estimation of a simple linear relation between the two censuses whose intercept incorporates any variation in coverage. It is also possible to estimate a slope that indicates the degree of registered death coverage from the average coverage of both censuses.

The SEG method (Bennett and Horiuchi, 1981), on the other hand, uses specific growth rates by age to convert a death age distribution into a population age distribution. More specifically, once the observed deaths at a given age x in the population are equal to the number of those aged x in the population, adjusted by the population growth rate by age range, the deaths at age $x+$ in a population provide an estimate of the number aged x in that population. The ratio of population above age x (estimated by reported deaths) to observed population above age x then represents the death registration coverage

The adjusted version of SEG is designed to deal with SEG’s highly problematic assumption that a population is closed to migration or has a very small migration flow, which is especially unrealistic in the case of Brazil and its regions. For example, when the population is not closed (migration flows) or when the two censuses give differential coverage, Hill et al. (2009) consider the GGB more robust than the SEG in simulations with age declaration errors in the census and death records. In simulations where census (death record) coverage varies (increases) by age, however, the GGB’s sensitivity often results in overestimation of the death registration coverage (Hill et al. 2009). The SEG method, in contrast, although relatively robust in the presence of age declaration errors in census and death records or when the difference in census coverage varies by age, produces substantially biased estimates in the presence of migration.

Although the literature proposes several methodologies for addressing these problems (Hill and Queiroz, 2010; Bhat, 2002), the adjusted SEG offers the simpler alternative of considering only age groups that are not greatly influenced by migration flows. That is, above a certain age x over which net migration is negligible, method performance minimizes the impact of migration flows. Estimate quality is further enhanced by using GGB-adj to combine the two methods (Hill, 2017), which Hill et al. (2009) suggest may be more robust than separate application. Because the most appropriate age interval selection technique for estimating underregistration involves assessing GGB-produced diagnostic charts, the adjusted method first applies GGB to obtain estimates of the change in population enumeration (k_1/k_2), which it uses to adjust the coverage of both censuses, and then applies SEG using the adjusted population for mortality data coverage. Given minimal differences in the final estimates, we adopt this combined method and use the DDM R-package to estimate the

undercounting of deaths for all 137 Brazilian mesoregions in the 1980–2010 period (<https://cran.r-project.org/package=DDM>).

Spatiotemporal analyses

To study the spatial evolution of adult mortality, we first use cluster analyses based on latent class models to determine spatial clustering across mesoregions and then employ spatial autocorrelation measures to determine the hotspots and outliers of adult mortality across the country. The latent class cluster (LCC) group model falls into the family of mixture likelihood clustering (McLachlan and Basford 1988; Everitt 1993), model-based clustering (Banfield and Raftery 1993; Bensmail et al. 1997; Fraley and Raftery 1998), mixture-model clustering (Jorgensen and Hunt 1996; McLachlan et al. 1999), Bayesian classification (Cheeseman and Stutz 1995), unsupervised learning (McLachlan and Peel 1996), and LCC analysis (Vermunt and Magidson, 2000). The method differs from standard cluster analysis techniques in its use of a model-based clustering approach to group information (Vermunt and Magidson 2002). That is, whereas classical cluster methods like k-means are commonly based on intuitively reasonable procedures (Everitt et al. 2011) but have difficulty deciding on the best distance method and estimation procedure for number of clusters, LCC applies statistical models for the population under study. In practical terms, LCC assumes that the population consists of many subpopulations or clusters with different multivariate probability density functions, defined by a finite mixture of underlying probability distributions (Everitt et al. 2011; Vermunt and Magidson 2002).

This finite mixture modeling is a form of latent variable analysis in which each subpopulation is a latent categorical variable and the latent classes are described by the different components of the mixture density. The allocation of objects to clusters is optimal when based on specific criteria that minimize the within-cluster variation and/or maximize the between-cluster variation (Vermunt and Magidson 2002). Whereas use of a statistical model offers the advantage of less arbitrary choice of cluster criterion, LCC allows great flexibility in the use of either simple or complicated distributional forms of the observed variables within clusters (Everitt et al. 2011; Vermunt and Magidson, 2002).

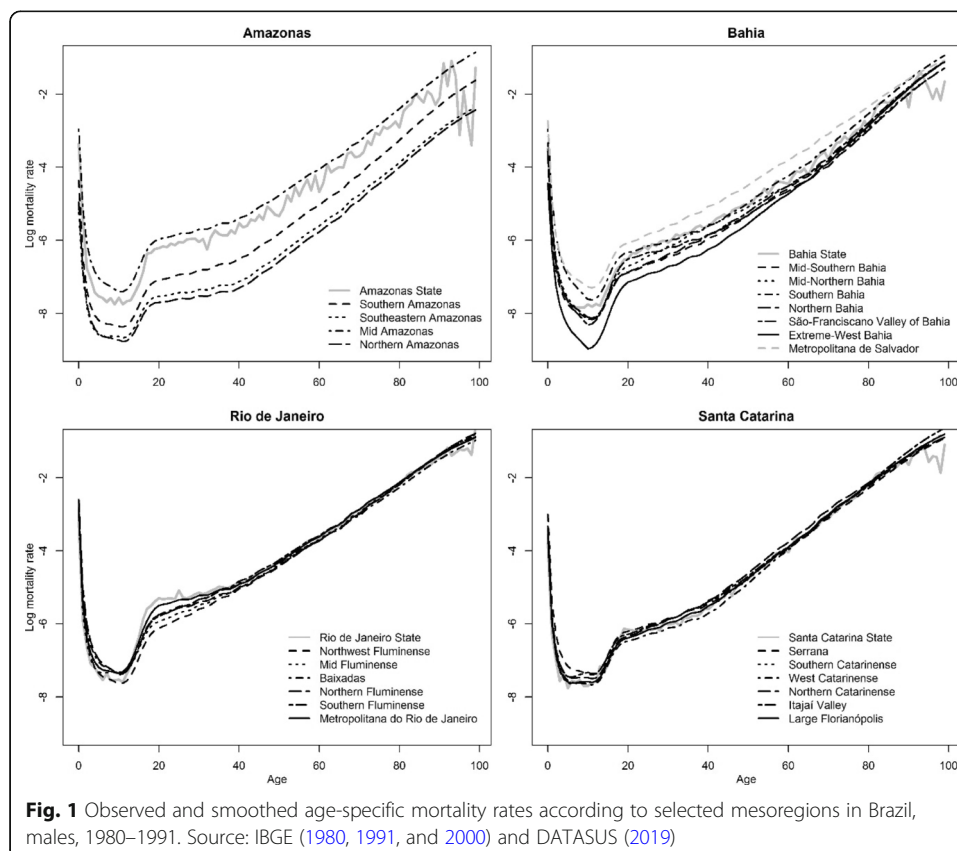
In our subsequent analysis of the spatial pattern of variation in adult mortality across Brazil's mesoregions, we use the Moran Index, a correlation coefficient measuring the linear relation between same variable values across neighboring areas. It is thus assigned the prefix “auto” and considered a spatial autocorrelation index (Bailey and Gatrell, 1995; Bivand et al., 2008), which may be global or local (Ward and Gleditsch, 2007; Anselin, 1995). Using global autocorrelation permits determination of whether the global spatial pattern is random (Anselin, 1995)—for example, whether the variation in adult mortality probability is higher in some subnational areas with a well-defined spatial gradient between them—while local autocorrelation enables measurement of the relation between closest neighboring areas. Certain situations require simultaneous analysis of both the local and global indexes: for example, using local autocorrelation to identify local spatial clusters or hotspots while assessing the influence of individual locations on the magnitude of the global statistic and identifying outliers (Anselin, 1993). The test of significance for these findings is based on pseudo-distribution references generated from simulations (Bailey and Gatrell, 1995; Bivand et al. 2008).

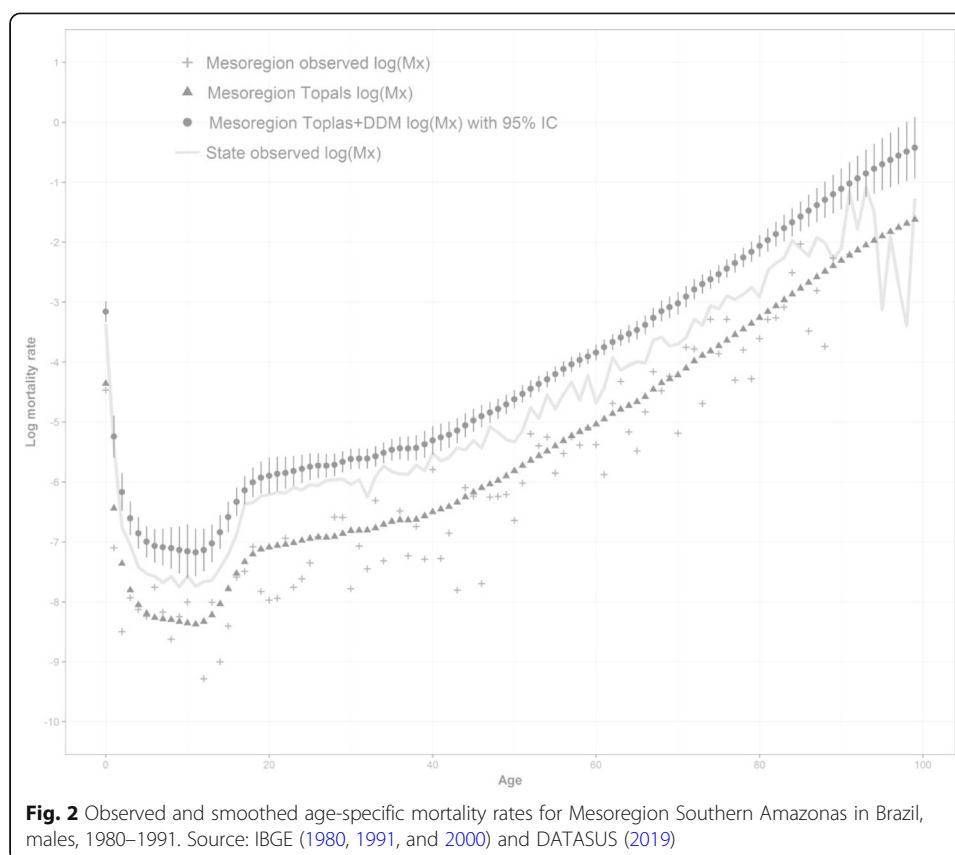
Results

Evolution of death count and adult mortality coverage completeness in Brazil

The distribution of smoothed male mortality rates across mesoregions in four different Brazilian states reveals substantial diversity in death data quality, from very poor in Amazonas and Bahia to relatively better in Rio de Janeiro and Santa Catarina (see Fig. 1). By plotting the mortality rate schedule for the entire state in each panel, we illustrate that mortality rates between mesoregions within some states may have different age schedules, especially at young adult ages, which could affect death probability between 15 and 60 years. Even on a state level, the ratio of deaths to exposition has higher variability across ages. In terms of mortality rate differences between mesoregions, however, whereas Rio de Janeiro and Santa Catarina show none, Amazonas and Bahia differ greatly, possibly as a result of different death record completeness. In general, the application of TOPALS yields reasonable age schedules for mortality rates in small areas, enabling accurate analysis of spatial and temporal variation across the country.

After smoothing mortality rates, we first apply DDM to estimate death count completeness and adjust mortality estimates when necessary and then more closely examine the estimates of smoothed plus adjusted (i.e., TOPALS + DDM) deaths for one select mesoregion, Southern Amazona (see Fig. 2). We denote the region's observed log mortality rates by +; its TOPALS estimates by small triangles; its TOPALS + DMM estimates, their 95% confidence intervals, and crossing segments by solid dots; and the observed log mortality rates for the entire state by a solid grey line. The most noteworthy observation





is the level of adjustment necessary to account for death underregistration in this mesoregion, which in fact reflects the tendency of most small areas in Brazil's Northern and Southern states to have the lowest death record completeness in the 1980–1991 period.

In Tables 1 and 2, by listing the descriptive statistics on coverage completeness for 1980–2010 death counts and adult mortality ($_{45}q_{15}$) in Brazil's five major regions for females and males (Tables 1 and 2), respectively, we show that the quality of national mortality data improved considerably in the later decades, with much less regional variation. For example, states in the South and Southeast recorded 100% of deaths in the most recent year, for both sexes, while some states in the North and Northeast, despite lower information quality for some areas, also show significant advances compared to earlier periods. As a result, by 2010, most mesoregions in the south-Southeast, as well as some in the Northeast and Midwest, had complete death registry coverage. Great progress is also evident in the quality of mortality information in the poorest mesoregions of the north-Northeast, especially those with the worst record quality in previous periods. In fact, based on the 95% confidence interval for $_{45}q_{15}$ estimates, these records reveal a significant decline in adult mortality for both sexes during the three decades examined, underscoring the summary statistics' usefulness in providing an overview of general changes across the country.

The estimates in Fig. 3, which are based on the TOPALS plus combined DDM (Hill et al. 2009), outline the spatial evolution of death count completeness by sex and mesoregion from 1980 to 2010. As the figure shows, during the 1980–1991 period, the Brazilian regions with the highest and most reliable data quality were those in the southern part of

Table 1 Completeness of Death Counts and Adult Mortality Rates for Females: Brazil and regions, 1980–2010

| Brazil | Completeness | | | 45q15 | | |
|------------------|------------------|------------------|------------------|---------------------------|---------------------------|---------------------------|
| | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .9497 | .9696 | .9367 | .135 (.126; .148) | .115 (.107; .125) | .108 (.101; .116) |
| Minimum | .1163 | .1299 | .3879 | .067 (.058; .079) | .063 (.051; .077) | .071 (.056; .089) |
| Maximum | 1 | 1 | 1 | .379 (.237; .570) | .288 (.181; .441) | .144 (.135; .204) |
| Total of areas | 137 | 137 | 137 | 137 | 137 | 137 |
| North | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .7692 | .8934 | .8528 | .154 (.126; .190) | .112 (.095; .132) | .109 (.097; .122) |
| Minimum | .1163 | .1299 | .3879 | .067 (.058; .079) | .063 (.052; .077) | .071 (.056; .089) |
| Maximum | 1 | 1 | 1 | .379 (.238; .571) | .288 (.181; .441) | .143 (.128; .204) |
| Total areas | 20 | 20 | 20 | 20 | 20 | 20 |
| Northeast | 1980–1991 | 1991–2000 | 2000–10 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .8375 | .9132 | .8864 | .112 (.104; .122) | .104 (.097; .111) | .106 (.099; .113) |
| Minimum | .2945 | .4922 | .5610 | .075 (.066; .085) | .076 (.070; .082) | .086 (.078; .093) |
| Maximum | 1 | 1 | 1 | .202 (.194; .210) | .159 (.153; .166) | .144 (.135; .154) |
| Total areas | 42 | 42 | 42 | 42 | 42 | 42 |
| Midwest | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .9478 | 1 | .9738 | .134 (.122; .147) | .119 (.110; .129) | .123 (.105; .121) |
| Minimum | .5182 | .8870 | .8658 | .076 (.061; .094) | .098 (.087; .110) | .094 (.091; .098) |
| Maximum | 1 | 1 | 1 | .157 (.151; .171) | .138 (.127; .153) | .129 (.120; .143) |
| Total areas | 15 | 15 | 15 | 15 | 15 | 15 |
| South | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | 1 | 1 | 1 | .137 (.129; .145) | .119 (.112; .126) | .104 (.099; .110) |
| Minimum | .9754 | .9537 | .9202 | .110 (.103; .119) | .100 (.093; .104) | .185 (.081; .090) |
| Maximum | 1 | 1 | 1 | .176 (.167; .187) | .148 (.140; .158) | .130 (.123; .138) |
| Total areas | 23 | 23 | 23 | 23 | 23 | 23 |
| Southeast | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | 1 | 1 | .9787 | .150 (.142; .159) | .127 (.120; .134) | .111 (.105; .117) |
| Minimum | .9523 | .8613 | .8295 | .122 (.116; .127) | .105 (.100; .110) | .093 (.087; .096) |
| Maximum | 1 | 1 | 1 | .179 (.171; .191) | .163 (.152; .175) | .140 (.132; .149) |
| Total areas | 37 | 37 | 37 | 37 | 37 | 37 |

Source: IBGE (1980, 1991, and 2000) and DATASUS (2019)

the country, with most northern areas characterized by very poor data quality. By 2010, however, most of Brazil's regions had completeness levels above 70% with far less variation in the best and worst completeness levels in each region. The estimates also reveal clear improvement in data quality in the Northeast regions over the 2000–2010 period, although the areas closest to state capitals² enjoy good coverage over the entire decade while others still need improvement over the next few years. Over the entire study period,

²In Brazil, all state capitals in the Northeast region are located in seacoast areas.

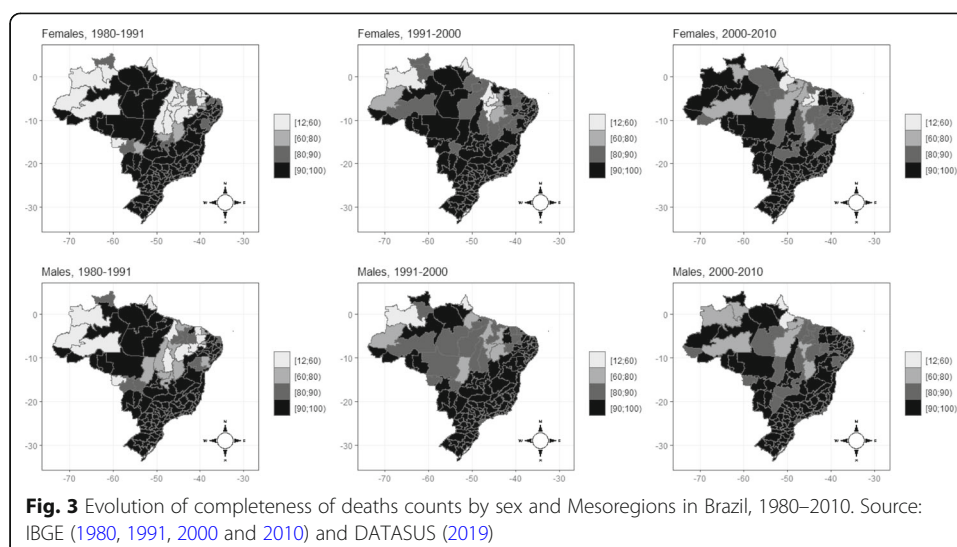
Table 2 Completeness of Death Counts and Adult Mortality Rates, Males, Brazil, and regions, 1980–2010

| | Completeness | | | 45q15 | | |
|------------------|------------------|------------------|------------------|---------------------------|---------------------------|---------------------------|
| | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Brazil | | | | | | |
| Mean | .9799 | .9639 | .9459 | .229 (.217; .244) | .216 (.205; .227) | .210 (.201; .220) |
| Minimum | .1804 | .2927 | .5875 | .097 (.086; .110) | .107 (.093; .122) | .133 (.121; .144) |
| Maximum | 1 | 1 | 1 | .445 (.331; .591) | .336 (.322; .433) | .289 (.277; .302) |
| Total areas | 137 | 137 | 137 | 137 | 137 | 137 |
| North | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .8818 | .8737 | .8609 | .223 (.197; .253) | .195 (.177; .216) | .199 (.185; .215) |
| Minimum | .1804 | .2927 | .5875 | .097 (.086; .110) | .107 (.093; .122) | .133 (.121; .144) |
| Maximum | 1 | 1 | 1 | .445 (.323; .591) | .332 (.281; .433) | .269 (.259; .301) |
| Total areas | 20 | 20 | 20 | 20 | 20 | 20 |
| Northeast | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .8922 | .9417 | .9276 | .188 (.177; .199) | .186 (.177; .196) | .201 (.192; .210) |
| Minimum | .3976 | .6882 | .6939 | .113 (.101; .125) | .138 (.127; .146) | .159 (.149; .170) |
| Maximum | 1 | 1 | 1 | .323 (.314; .332) | .303 (.298; .308) | .275 (.267; .282) |
| Total areas | 42 | 42 | 42 | 42 | 42 | 42 |
| Midwest | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | .9256 | .9844 | .9517 | .236 (.222; .252) | .227 (.215; .239) | .222 (.212; .233) |
| Minimum | .4593 | .7961 | .8225 | .160 (.139; .184) | .185 (.170; .201) | .191 (.179; .202) |
| Maximum | 1 | 1 | 1 | .285 (.269; .307) | .250 (.235; .266) | .257 (.248; .269) |
| Total areas | 15 | 15 | 15 | 15 | 15 | 15 |
| South | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | 1. | 1. | 1. | .240 (.230; .250) | .223 (.214; .232) | .206 (.199; .214) |
| Minimum | .9733 | .9531 | .9147 | .196 (.188; .204) | .183 (.174; .190) | .168 (.162; .174) |
| Maximum | 1 | 1 | 1 | .280 (.274; .286) | .260 (.255; .265) | .237 (.227; .247) |
| Total areas | 23 | 23 | 23 | 23 | 23 | 23 |
| Southeast | 1980–1991 | 1991–2000 | 2000–2010 | 1980–1991 (95% IC) | 1991–2000 (95% IC) | 2000–2010 (95% IC) |
| Mean | 1. | .9897 | .9727 | .271 (.261; .282) | .251 (.242; .260) | .224 (.216; .232) |
| Minimum | .9335 | .8472 | .8182 | .219 (.210; .228) | .207 (.197; .217) | .191 (.186; .197) |
| Maximum | 1 | 1 | 1 | .336 (.331; .352) | .336 (.322; .351) | .289 (.277; .302) |
| Total areas | 37 | 37 | 37 | 37 | 37 | 37 |

Source: IBGE (1980, 1991, and 2000) and DATASUS (2019)

death count coverage is higher for males than for females, especially in North Brazil, although the sex difference does seem to decrease during the 2000–2010 decade.

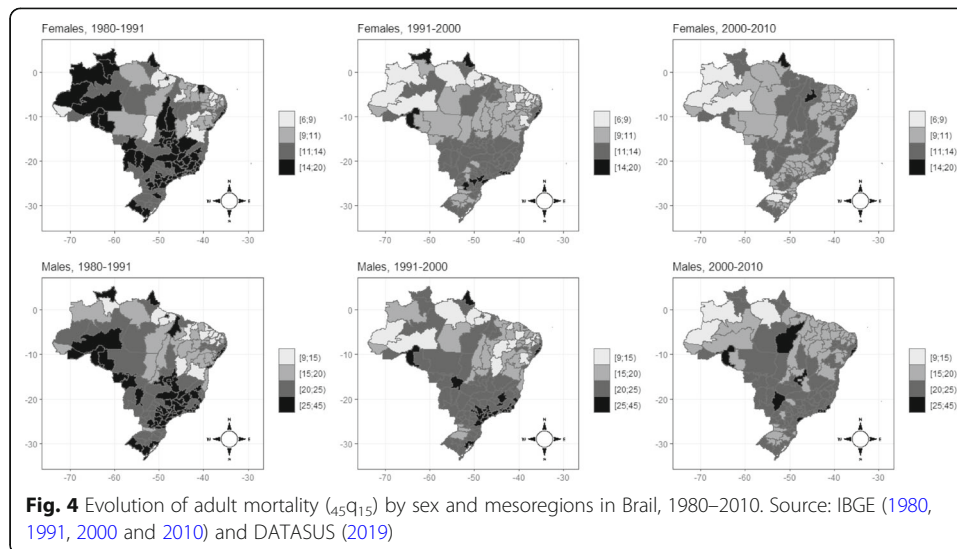
Two possible reasons for this decline are a major increase during this period in mortality from external cause, especially among males, and better registration of external death causes (Queiroz, et al. 2017; Moura, et al. 2015; Lima and Queiroz, 2014; Waiselfiz, 2013). In the literature, there is evidence in Brazil that external causes of death (such as homicides and transit accidents) have almost complete coverage in all Brazilian regions (Murray, et al. 2013; Campos and Rodrigues 2004). As summarized by



Schmertmann and Gonzaga (2018) violent causes of deaths due to homicides and traffic accidents must be reported to the local health department and to the police. Hence, there are more sources of information for this particular death cause, improving its quality in turn. One may also argue that there is variation in the quality of this information across regions of the country, but the process of registering this type of death is defined by a federal law (Murray, et al. 2013). In addition, external causes of deaths, especially for males, have increased in the last decades, and these deaths are concentrated among young male adults (Murray, et al. 2013; Soares Filho et al., 2016; França, et al. 2017; Borges, 2017).

In relation to the regional pattern of female mortality, one should highlight that the country has not fulfilled its commitment to reduce 75% of maternal deaths until 2015. In 2015, the ratio of maternal mortality was around 62 deaths per 100,000 live births. There are also large regional inequalities, regions in the North and Northeast parts of the country present much higher maternal mortality than the regions in the Southern parts of the country (Silva et al., 2016). Rodrigues, et al. (2016) shows an increase in maternal mortality in regions in the North and Northeast and a steady decline in the Southeast and South. The authors also suggest that some areas have shown an increase in maternal mortality in the 2000s. Despite this rising trend in maternal deaths, as suggested by Borges (2017), for larger areas, in recent decades it is observed an increased gap between male and female life expectancy. The main explanation for that is the elevated male mortality by external causes of death. Baptista and Queiroz (2019) also show that mortality by cardiovascular diseases is lower and declines faster for females than for males.

While death count coverage increased over the three decades, however, adult mortality for both sexes and all mesoregions declined, as shown by the evolution of death probability at age 15–60 ($_{45}q_{15}$) across the 1980–2010 period (see Fig. 4). Nonetheless, despite a seemingly similar pattern of decline for males and females, male mortality remained higher in some parts of the country, especially in the Southeast, coastal Northeast, and a few Northern areas where the risk of dying aged 15–60 was around 0.25 for males. One possible explanation is a rise in violence rates (Moura et al., 2015; Waiselfiz, 2013;

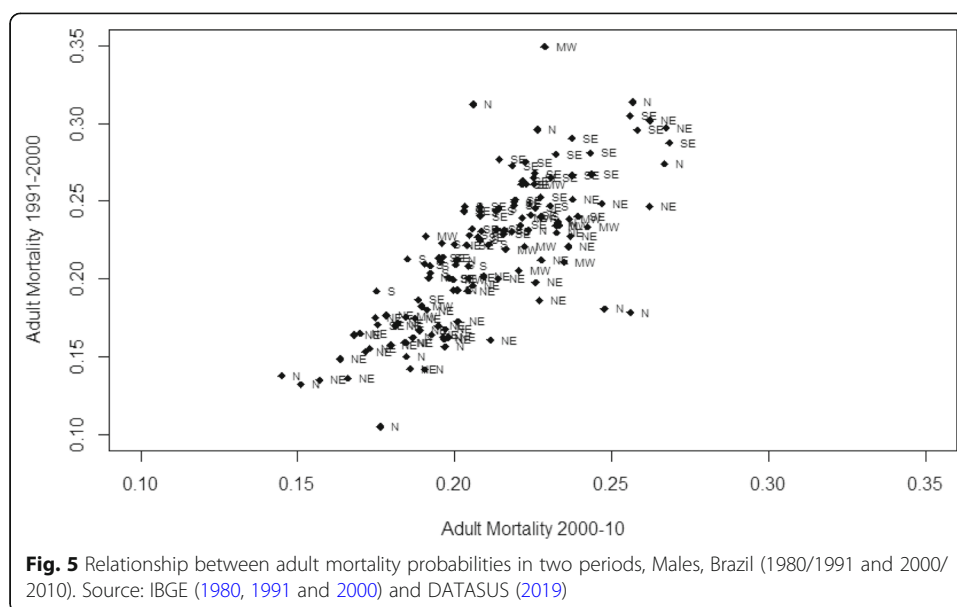


Reichenheim et al. 2011), which although they decreased in the Southeast between 1990 and 2010, remained high, with a relative increase in the Northeast during the same period (Waiselfiz, 2013; de Andrade and Diniz, 2013; Carvalho et al. 2012; Pereira and Queiroz, 2016). At the same time, the 95% confidence interval signals that the variation in adult mortality reduced over time and across regions, indicating a quickly declining level of uncertainty, whose higher (lower) level in less (more) populated and less (more) developed areas of the country reflects the improvement in data quality. Most Northern and North-eastern areas also show a slowing pace of mortality decline for both sexes, which may be attributable to two factors: the increased risk of death from external causes in these regions (Borges, 2017; Waiselfiz, 2013), and a large decline in mortality level over the 1990–2000 decade that set a low starting point for levels in the next decade.

In Fig. 5, by plotting the death probability at ages 15–60 for both sexes in 1980–1991 against that in 2000–2010, with the $_{45}q_{15}$ estimates for the former (latter) on the x -axis (y -axis), we demonstrate that the regions with a high death probability in the 1980s are still those with high mortality levels in the later periods. Hence, although adult mortality ($_{45}q_{15}$) differs substantially across Brazilian regions over the three decades of study, most regions with high mortality levels in 1980–1991 also had high levels in 2000–2010, with similar trends observable for both males and females.

Spatial analysis of adult mortality patterns

The process of divergence and slow convergence in Brazil's adult mortality in recent years reflects the slow rhythm of mortality decline in some regions of the country (Borges, 2017; França, et al., 2017; Baptista and Queiroz, 2019). That is, in general, regions with high mortality levels in 1980 show a slower decline in adult mortality but have higher levels in the more recent periods. One way to examine this trend is to estimate—and then spatially analyze—the variation in intercensal adult mortality probabilities, which, in addition to further explaining earlier observations, should shine light on the corresponding mortality dynamics across time and space. To this end, we divide our three study decades into two periods and perform cluster estimates using centroids from the

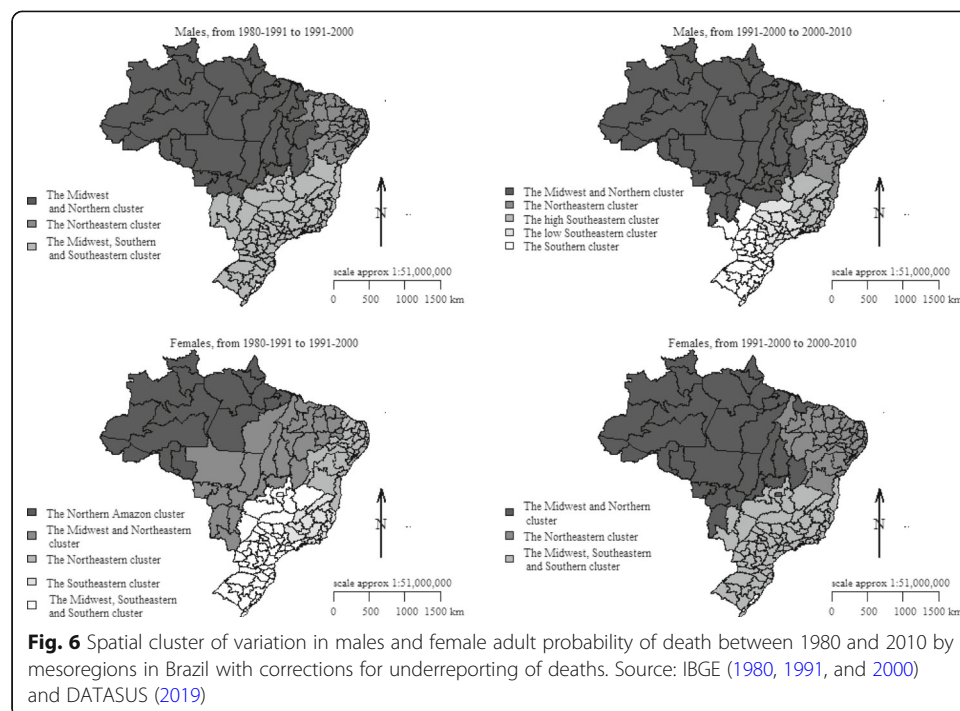


mesoregions as spatial control variables, together with the variation in death probability. Figure 6 shows spatial cluster of variation in adult mortality for two intercensal periods.

For males in study period 1 (1980–1991 to 1991–2000), we identify three spatial clusters, the first encompassing the entire North and parts of the Northeast and Midwest. This cluster comprises mainly mesoregions with adult mortality reduction (69.7%), unchanged variability (12.1%), and increased adult death probability in 18.2% areas over time. The second cluster, in the Northeast, is more heterogeneous and comprises mesoregions with no change in mortality risk (48.5%) during period 1, 30.3% of whose areas actually experienced a decline in adult mortality between 1980–1991 and 1991–2000. The third cluster is similar but includes no mesoregions with increased death probability across these two decades.

For males in period 2 (1991–2000 to 2000–2010), we identify five clusters of mortality variation, including the same Midwest and Northern groups as in period 1, while also noting that most areas show either no change (35.9%) or an increase (38.5%) in adult mortality. Only the Northeast cluster includes more areas with increased mortality risk (75%). Two additional groups of mesoregions in the Southeast, South, and parts of the Midwest are predominantly characterized by adult mortality reduction. The Southeast is divisible into two spatial clusters: the Upper Southeastern region, in which 74% of areas experience a mortality reduction but 26% show no significant change; and the Lower Southeastern cluster, in which 100% of the mesoregions experience mortality decline. The profile of the latter is nonetheless similar to that of the former, with 80% of areas experiencing a decrease and 20% showing no change in $_{45}q_{15}$ mortality. It is also worth noting that not all mesoregions show only an increase or only a decrease in death probability for any one cluster, prompting us to describe mortality levels based on the number of areas that experienced a reduced or increased mortality risk.

For females in period 1, we identify a fourth spatial cluster whose profile, with few exceptions, changes little compared to that of males. In fact, even though we find distinct spatial patterns of variation for five and three female clusters in the first and second



periods, respectively, these patterns differ only slightly from those of males. Across the time span, the Northern and high Midwest clusters are again the most heterogeneous, with mortality reduction being slightly predominant (33.3%) and 41% of the mesoregions experiencing increased adult mortality in period 2. The Northeastern spatial cluster is characterized mainly by increased mortality (53.3% of areas), with 37.8% showing no change in $_{45}q_{15}$ variation. Both these clusters show a similar pattern of adult mortality variation, with a majority of mesoregions experiencing decreasing adult mortality.

The above cluster analysis of Brazilian mesoregions across our two study time periods (1980–1991 to 1991–2000; 1991–2000 to 2000–2010) reveals two particularly interesting patterns: First, during the first 20 years of adult mortality change in the country, although some clusters in the Northeast and some parts of the North experienced increased mortality risk, levels in Brazil's most developed areas barely changed. Second, although this pattern of increased mortality risk in the Northeast and some Northern areas persisted over the entire time span, this risk either decreased or remained unchanged in the most developed regions, especially the South and Southeast. This geographic difference in profile strongly suggests an association between lower mortality levels and socioeconomic development.

Spatial analyses of adult mortality variation in Brazil

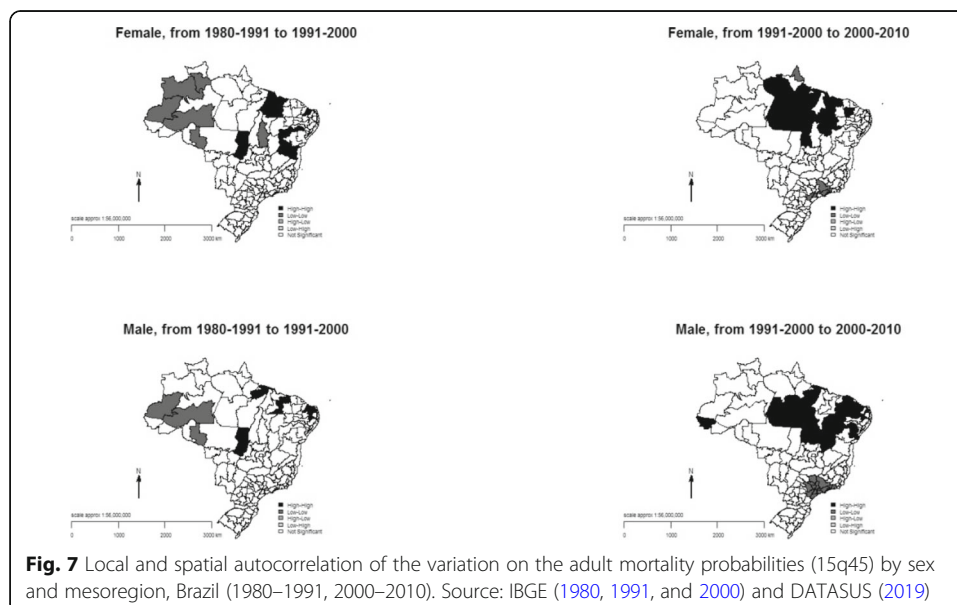
Our last investigative phase is a spatial auto-correlation analysis using global and local Moran Indexes to measure adult mortality variation by sex and region between period 1 (1980–1991 to 1991–2000) and period 2 (1991–2000 to 2000–2010). To assess the presence of spatial clusters of high and low change in adult mortality ($_{45}q_{15}$), we employ the local Moran's I statistic as a local indicator of spatial association (LISA; Anselin 1995; Cliff and Ord 1975). Our specific focus is whether relative reductions in adult mortality are randomly distributed or concentrated in particular areas. The significant

positive values on the global index for both males and females indicate that the spatial variation in $_{45}q_{15}$ is not random.

In Fig. 7, the areas designated high-high (low-low) are those that experienced a high (low) variation in mortality in period 1 relative to period 2, with the outliers, high-low and low-high, being regions with significant heterogeneity in mortality variation but a unique pattern of either high or low variation in each specific area. None of these outliers are significant, however, at a 0.05 significance level. In period 1, some regions in the North are characterized by low-low areas, suggesting low adult mortality variation across the northern region and surrounding areas, except for a small group of northern regions in the Northeast, which show predominantly high-high variation. This latter indicates neighboring mesoregions with high adult mortality variation for both men and women. Overall, during Period 1, the regions with significant adult mortality variation—both low-low and high-high—are located in northern Brazil.

In period 2, the areas of homogeneously high variation in adult mortality, whose spatiotemporal pattern holds for both males and females, are a set of mesoregions in the north central portion of the Northeast, semi-arid mesoregions in the Northeast, and mesoregions in the North. As previously explained, because very few Brazilian regions show an increase in mortality during the analytic period, we can treat this variation as a decline in adult mortality. The results further indicate that in much of northern Brazil, adult mortality reduction is low in several neighboring areas in period 1 and accentuated in several mesoregions Period 2. A similar pattern is observable in the north central part of the Northeast, which, compared with high-high areas in the two different periods, shows some indication of a high-high concentration, possibly implying contagion on the rate of adult mortality decline over time in that area. The south central portion of the country, in contrast, is characterized by a low-low spatial cluster in period 2 that is concentrated in the Southeast, a result that is hardly surprising given that most Southeast states already had the lowest mortality rates in the country.

Despite the overall decline in adult mortality levels over time, however, we note an evolution of $_{45}q_{15}$ homogenization and convergence across the country. That is, even



with the persistence of hotspots in certain locations, Brazilian adult mortality levels are becoming more alike across mesoregions, especially among females. On the other hand, our examination of adult mortality convergence across Brazilian mesoregions between 1980 and 2010 reveals significant differences, as well as high mortality hotspots in some areas as a result of increasing deaths from external causes, especially among young males (Waiselfiz, 2013; Pereira and Queiroz, 2016). At the same time, reduced mortality rates in other areas with lower socioeconomic levels may imply better living conditions and overall improvements in health.

Conclusions

The empirical findings reported in this paper make a valuable contribution not only to the research on mortality in general but also to that on data quality and adult mortality in small areas of Brazil, whose accuracy tends to be hampered by defective data. Our combined TOPALS and DDM methodology, in addition to assessing the quality of Brazil's regional death count information, produces accurate adult mortality estimates (and confidence intervals) for these small areas by age, sex, and population in a manner easily replicable in all countries that suffer from data deficiencies (Tabutin et al., 2017). Nonetheless, although this method is applicable to all analytic periods and geographic areas—a major advantage of our methodological approach—it suffers from one major limitation; namely, the potential for population size and data quality-related errors when estimating data completeness and adult mortality rates for small areas. Although, a detailed discussion of these errors is beyond the scope of this (and many other) papers (cf. Baker et al., 2013a, b), we do control for them by focusing on adult mortality among 5-year age groups in mesoregions, a unit far less volatile than city or census tract. We also provide 95% confidence intervals for the mortality age profiles, thereby designating the uncertainty level for each estimate across region and time. As a final precaution, we use regions with at least 53,934 inhabitants, and in additional unreported analyses compare our main results with outcomes using an empirical Bayesian estimator. The similarity of these findings supports our confidence in the robustness of our results.

The main observations from our study, which resemble those documented for larger areas (Paes, 2005; Paes, 2007; Lima and Queiroz, 2014; Queiroz, et al. 2017), are that the quality of death count registration in Brazil has been improving over time in all its regions while the adult mortality rate has been declining, albeit with a notable gap between males and females. On the other hand, not only are certain areas of the country still characterized by stagnant or increasing mortality levels because of violence-related young adult deaths, but our examination of the evolving completeness of death count and adult mortality coverage in Brazil reveals remarkable regional differences. For example, whereas states in the more developed South and Southeast achieved completeness levels close to 100% over time for both sexes; some states in less developed Northeast and Northern regions still have lower quality information but have made significant recent advances over the 1980s.

Similar improvement in death count completeness occurred over the study period in Brazil's North and Northeast, with areas closest to state capitals having higher coverage in all decades (Paes, 2005; Paes, 2007; Lima and Queiroz, 2014; Queiroz, et al. 2017). The converse of this pattern is confirmed by active search-based estimations of infant

death coverage in these areas (Szwarcwald et al. 2014), which suggest that underreporting is probably lower than official estimations (IBGE, 2013). By 2010, therefore, most states in south-Southeastern regions, as well as some in Northeast and Midwest regions, had complete death registry coverage, and much progress had been made in improving mortality information quality even in the poorest north-Northeastern areas, particularly those with the worst record quality in the earlier decades.

In Brazil, there are few studies that applied spatial and demographic analyses on overall mortality (Baptista and Queiroz, 2019; Schmertmann and Gonzaga, 2018). In general, the focus is on epidemiological and public health aspects of mortality, and the studies' target are the entire country, major geographic regions (Southeast, Northeast, Federal States, etc.), or very specific counties. Therefore, there is a need for studies investigating the variation of mortality across time and space using smaller areas. In this paper, we show that the epidemiologic transition in Brazil has not followed the linear and unidirectional pattern, and it is quite heterogeneous across regions of Brazil and within regions. Moreover, we argue that additional studies should pay attention to social and regional inequalities in mortality. In this sense, this study helps to understand the dynamics of the health transition in the country over a long period of time, and it may have positive impacts on local public health policies. Our results are in line with other analyses that focus on states or major regions of the country, but we provide more detailed results for small regions.

Although the adult mortality estimates suggest improved health conditions in the country over the study period, the differences in adult mortality by sex have remained almost constant, primarily because of male deaths from external causes (Pereira and Queiroz, 2016; Malta, et al. 2017; Ladeira, et al. 2017). Such deaths also partly explain the very small variation in 45q15 across regions over time: above 0.200 for males; around 0.120 for females, with intersex differences remaining practically constant from 2000 to 2010. Not only is female adult mortality far lower than male mortality, with the highest death probability for 2010 found in Rio de Janeiro, Espírito Santo, Alagoas, and Pernambuco, but the regions with the fastest decline in adult mortality are those with highest mortality levels in earlier decades. For instance, the burden of disease is higher in the North and Northeast, while chronic diseases like cardiovascular, diabetes, and obstructive pulmonary disease predominate in all other regions (Souza et al., 2017). Mental disorders, especially depression, homicide, and traffic accidents are also higher in the region and present a huge public health problem (Borges, 2017; Waiselfiz, 2013).

The remarkable changes in adult mortality across Brazilian regions over the three decades studied are particularly underscored by the spatio-temporal cluster analysis whose spatial statistics of death probability variation reveal a homogeneous increase in adult mortality in the North and most of the Northeast between 1980 and 1991. Likewise, the increase in adult mortality during the first study decade (2000–2010) was more persistent in the Northeast, mainly for males, probably because between 2002 and 2010, violence in this region increased 12.2 % for whites and 96.7% for blacks compared to respective decreases of 50.8% and 30.1% in the Southeast (Waiselfiz, 2013). In fact, the South and Southeast saw a persistent decline in adult mortality for both sexes in all three decades of our analysis.

These patterns of increasing adult mortality in the north-Northeast and reductions in the south-Southeast are consistent with changes in the age pattern for cause of death, suggesting that central and local government efforts to improve Brazil's data quality

(Borges, 2017; Frias, et al. 2017) are bearing fruit. Yet although the improved data may enhance understanding of the dynamics of health and mortality transitions in the country, more investment is clearly necessary in areas that still lag behind in such upgrading. In particular, there must be continuous study and evaluation of data quality, especially for small areas, and investment by all administrative levels into improving health information in Brazil (Mello Jorge et al., 2010).

Indeed, the significant improvement in the quality of adult mortality data observed in our study appears related to investment in the public health care system and administrative procedures to improve the recording of vital events. Hence, attempts to improve data quality may be significantly positively impacted by investment in such initiatives as the Family Health Program, which works closely with the community and monitors the health status of several individuals in each location. Future research might also build upon results reported in this paper to perform more in-depth analyses of data quality on a state level, provide substantive empirical evidence of the mortality differential between males and females, and expand understanding of the social and economic determinants of Brazil's mortality differential overall.

Acknowledgements

We would like to thank several seminars participants (EPC, PAA, ABEP) at different events where previous versions of the paper were presented. Julia Angelica for revising and editing the original English version. We thank the two anonymous reviewers whose comments/suggestions helped improve and clarify this manuscript. All data and codes are available at https://demografiaufrn.net/laboratorios/lepp/paper_genus/.

Authors' contributions

BLQ: conceptualization, methodology, validation, formal analysis, investigation, writing–review and editing, visualization; EECL: conceptualization, methodology, formal analysis, investigation, writing–review and editing visualization; FHF: conceptualization, methodology, formal analysis, investigation, writing–review and editing; MRG: conceptualization, methodology, formal analysis, investigation, writing–review and editing. All authors read and approved the final manuscript.

Funding

This work is part of a joint project for estimating mortality and constructing life tables for small regions in Brazil (1980–2010) by the Ministry of Science, Technology, Innovation, and Communications (MCTI); the National Council for Scientific and Technological Development (CNPq); the Ministry of Education (MEC); the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES); Applied Social Sciences (Process No. 470866/2014-4); and MCTI/CNPq/Universal 14/2014 (Process No. 454223/2014-5). Universidade Federal do Rio Grande do Norte – (Process No. 23077.001539/2020-40)

Availability of data and materials

The datasets analyzed during the current study are public available in the Ministry of Health Database (www.datasus.gov.br) and Brazilian Census Bureau (www.ibge.gov.br). All data and codes are available at https://demografiaufrn.net/laboratorios/lepp/paper_genus/.

Competing interests

All authors declare no competing interests.

Author details

¹Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627–Cedeplar/Face/UFMG, Belo Horizonte 31270-901, Brazil. ²Universidade Estadual de Campinas, Campinas, Brazil. ³Universidade Federal do Rio Grande do Norte, Natal, Brazil.

Received: 24 April 2020 Accepted: 1 October 2020

Published online: 02 November 2020

References

- Adair, T., & Lopez, A. D. (2018). Estimating the completeness of death registration: an empirical method. *PLoS one*, 13(5).
- Ahmed S, Hill K. (2011). Maternal mortality estimation at the subnational level: a model-based method with an application to Bangladesh. *Bulletin of the World Health Organization*, 89(1):12–21.
- Alexander, M., Zagheni, E., & Barbieri, M. (2017). A flexible Bayesian model for estimating subnational mortality. *Demography*, 54(6), 2025–2041.
- Anselin, L. (1993). *The Moran scatterplot as an ESDA tool to assess local instability in spatial association*. Morgantown, WV: Regional Research Institute, West Virginia University.
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographic Anal*, 27(2), 93–115.

- Assunção, R. M., Schmertmann, C. P., Potter, J. E., & Cavenaghi, S. M. (2005). Empirical Bayes estimation of demographic schedules for small areas. *Demography*, 42(3), 537–558.
- Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*, (vol. 413). Essex: Longman Scientific & Technical.
- Baker, J., Alcantara, A., Ruan, X., Watkins, K., & Vasan, S. (2013b). A comparative evaluation of error and bias in census tract-level age/sex-specific population estimates: component I (net-migration) vs component III (Hamilton–Perry). *Population Research and Policy Review*, 32(6), 919–942.
- Baker, J. D., Alcantara, A., Ruan, X., Vasan, S., & Nathan, C. (2013a). An evaluation of the accuracy of small-area demographic estimates of population at risk and its effect on prevalence statistics. *Popul Health Metrics*, 11(1), 24.
- Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 803–821.
- Baptista, E. A., & Queiroz, B. L. (2019). Spatial analysis of mortality by cardiovascular disease in the adult population: a study for Brazilian micro-regions between 1996 and 2015. *Spatial Demography*, 7(1), 83–101.
- Barufi, A. M., Haddad, E., & Paes, A. (2012). Infant mortality in Brazil, 1980–2000: A spatial panel data analysis. *BMC Public Health*, 12(1), 181.
- Bennett, N. G., & Horiuchi, S. (1981). Estimating the completeness of death registration in a closed population. *Popul Index*, 207–221.
- Bensmail, H., Celeux, G., Raftery, A. E., & Robert, C. P. (1997). Inference in model-based cluster analysis. *Stat Comput*, 7(1), 1–10.
- Bernardinelli, L., & Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat Med*, 11(8), 983–1007.
- Bhat, P. N. M. (2002). General growth balance method: a reformulation for populations open to migration. *Popul Stud*, 56(1):23–34.
- Bivand, R. S., Pebesma, E. J., Gomez-Rubio, V., & Pebesma, E. J. (2008). *Applied spatial data analysis with R*, (vol. 747248717). New York: Springer.
- Borges, G. M. (2017). Health transition in Brazil: regional variations and divergence/convergence in mortality. *Cadernos SaúdePública*, 33(8).
- Campos, N. O. B., & Rodrigues, R. N. (2004). Ritmo de declínio nas taxas de mortalidade dos idosos nos estados do Sudeste, 1980–2000 [The pace of decline in mortality rates of the elderly in states of the Southeast, 1980–2000]. *Rev Bras Estudos Popul*, 21, 323–342.
- Carvalho, A. X. Y. D., Silva, G. D. M. D., Almeida Júnior, G. R. D., & Albuquerque, P. H. M. D. (2012). Taxas bayesianas para o mapeamento de homicídios nos municípios brasileiros. *CadSaude Publica*, 1249–1262.
- Cheeseman, P., & Stutz, J. (1995). Bayesian classification (AutoClass): Theory and results. In U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in knowledge discovery and data mining*, (pp. 153–180).
- Cliff, A. D., & Ord, J. K. (1975). Model building and the analysis of spatial pattern in human geography. *J Royal Stat Soc Series B (Methodological)*, 37(3), 297–348.
- DATASUS (2019). Estatísticas Vitais. Available at <http://www2.datasus.gov.br/DATASUS/index.php?area=0205>. Accessed 20 June 2020.
- de Andrade, L. T., & Diniz, A. M. A. (2013). A reorganização espacial dos homicídios no Brasil e a tese da interiorização. *Rev Bras Estudos Popul*, 30, 171–191.
- De Beer, J. (2012). Smoothing and projecting age-specific probabilities of death by TOPALS. *Demographic Res*, 27(20), 543–592.
- Divino, F., Egidi, V., & Salvatore, M. A. (2009). Geographical mortality patterns in Italy: A Bayesian analysis. *Demographic Res*, 20, 435–466.
- Dorrington, R. E. (2014a). General Growth Balance. In Moultrie TA, RE Dorrington, AG Hill, KH Hill, IM Timæus and B Zaba (eds), *Tools for Demographic Estimation*. <http://demographicestimation.iussp.org/content/general-growth-balance>.
- Dorrington, R. E. (2014b). Synthetic extinct generations. In Moultrie TA, RE Dorrington, AG Hill, KH Hill, IM Timæus and B Zaba (eds), *Tools for Demographic Estimation*. <http://demographicestimation.iussp.org/content/synthetic-extinct-generations>.
- Ehrl, P. (2017). Minimum comparable areas for the period 1872–2010: an aggregation of Brazilian municipalities. *Estudos Econ (São Paulo)*, 47(1), 215–229.
- Everitt, B. S. (1993). *Cluster analysis*. London: Edward Arnold.
- Everitt B S., Landau S, Leese M and Stahl D. (2011). *Cluster Analysis*, 5th Edition. Wiley Series in Probability and Statistics.
- Fenelon, A. (2013). Geographic divergence in mortality in the United States. *Popul Dev Rev*, 39(4), 611–634.
- Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering methods? Answers via model-based cluster analysis. *Comp J*, 41, 578–588.
- França, E., Abreu, D. M. X., Rao, C., & Lopez, A. D. (2008). Evaluation of cause of death statistics for Brazil, 2002–2004. *Int J Epidemiol*, 37(4), 891–901.
- França, E. B., de Azeredo Passos, V. M., Malta, D. C., Duncan, B. B., Ribeiro, A. L. P., Guimaraes, M. D., ... & Camargos, P. (2017). Cause-specific mortality for 249 causes in Brazil and states during 1990–2015: a systematic analysis for the global burden of disease study 2015. *Population health metrics*, 15(1), 1–17.
- Frias, P. G. D., Szwarcwald, C. L., MoraisNeto, O. L. D., Leal, M. D. C., Cortez-Escalante, J. J., Souza Junior, P. R. B. D., ... Silva Junior, J. B. D. (2017). Use of vital data to estimate mortality indicators in Brazil: from the active search for events to the development of methods. *Cadernos SaúdePública*, 33(3).
- Gonzaga, M. R., & Schmertmann, C. P. (2016). Estimating age-and sex-specific mortality rates for small areas with TOPALS regression: an application to Brazil in 2010. *Rev Brasil Estud Popul*, 33(3), 629–652.
- Hill, K. (1987). Estimating census and death registration completeness. *Asian Pac Census Forum*, 1(3), 8–13 23–24.
- Hill, K. (2017). Analytical methods to evaluate the completeness and quality of death registration: current state of knowledge. Population Division Technical Paper, no. 2017/02.
- Hill, K., Choi, Y., & Timæus, I. (2005). Unconventional approaches to mortality estimation. *Demographic Res*, 13, 281–300.
- Hill, K., You, D., & Choi, Y. (2009). Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors. *Demographic Res*, 21, 235–254.
- Hill K, Queiroz B. (2010). Adjusting the general growth balance method for migration. *Rev Bras Estud Popul*, 27(1):7–20.
- IBGE (2008). Metodologia das estimativas das Populações residentes nos municípios brasileiros para 1o de Julho de 2008: Uma abordagem demográfica para estimar o padrão histórico e os níveis de subenumeração de pessoas nos censos demográficos e contagens de população. IBGE, 30 p.

- IBGE (2013) Tábuas Abreviadas de mortalidade por sexo e idade: Brasil, Grandes Regiões e Unidades da Federação, 2010. Estudos e pesquisas. Informação Demográfica e Socioeconômica, n. 30. Rio de Janeiro: IBGE, 2013. Available at: http://ibge.gov.br/home/estatistica/populacao/tabuas_abreviadas_mortalidade/2010/default.shtm. Accessed 7 July 2019.
- IBGE, Instituto Brasileiro de Geografia e Estatística (1980). Censo demográfico: 1980: dados gerais, migração, instrução, fecundidade, mortalidade. IBGE, <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?id=772&view=detalhes>. Accessed 4 Nov 2018.
- IBGE, Instituto Brasileiro de Geografia e Estatística (1991). Censo demográfico: 1991: dados gerais, migração, instrução, fecundidade, mortalidade. IBGE, <https://biblioteca.ibge.gov.br/biblioteca-catalogo?id=782&view=detalhes>. Accessed 4 Nov 2018.
- IBGE, Instituto Brasileiro de Geografia e Estatística (2000) Censo Demográfico 2000. Amostra - Características Gerais da População. <https://sidra.ibge.gov.br/pesquisa/censo-demografico/demografico-2000/amostra-caracteristicas-gerais-da-populacao>. Accessed 4 Nov 2018.
- IBGE, Instituto Brasileiro de Geografia e Estatística (2010). Censo demográfico, 2010. Características da População e dos Domicílios. IBGE, http://www.ibge.gov.br/home/estatistica/populacao/censo2010/sinopse/default_sinopse.shtm. Accessed 4 Nov 2018.
- Jorgensen, M., and Hunt, L. (1996). Mixture model clustering of data sets with categorical and continuous variables, in Proceedings of the Conference ISIS, Vol. 96), 375–384.
- Kibele, E. U., Klüsener, S., & Scholz, R. D. (2015). Regional mortality disparities in Germany. *KZfSS Kölner Zeitschrift für Soziologie Sozialpsychologie*, 67(1), 241–270.
- Kulkarni, S. C., Levin-Rector, A., Ezzati, M., & Murray, C. J. (2011). Falling behind: life expectancy in US counties from 2000 to 2007 in an international context. *Popul Health Metrics*, 9(1), 16.
- Ladeira, R. M., Malta, D. C., Moraes Neto, O. L. D., Montenegro, M. D. M. S., Soares Filho, A. M., Vasconcelos, C. H., ... Naghavi, M. (2017). Road traffic accidents: Global Burden of Disease study, Brazil and federated units, 1990 and 2015. *Rev Bras Epidemiol*, 20, 157–170.
- Leone, T. (2014). Measuring Differential Maternal Mortality Using Census Data in Developing Countries. *Population, Space and Place*, 20(7):581–591
- Lima, E. E. C. D., & Queiroz, B. L. (2014). Evolution of the deaths registry system in Brazil: associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. *Cadernos Saúde Pública*, 30(8), 1721–1730.
- Lima, E. E. C. D., Queiroz, B. L., Missov Trifon & Lenart, Adam (2016). Methods to estimate mortality curves in small areas: an application to municipality data in Brazil. In: Population Association of America Annual Meeting, Washington.
- Luy, M. (2011). A classification of the nature of mortality data underlying the estimates for the 2004 and 2006 United Nations' World Population Prospects. *Comparative Population Studies*, 35(2).
- Malta, D. C., Minayo, M. C. D. S., Soares Filho, A. M., Silva, M. M. A. D., Montenegro, M. D. M. S., Ladeira, R. M., ... Naghavi, M. (2017). Mortality and years of life lost by interpersonal violence and self-harm: in Brazil and Brazilian states: analysis of the estimates of the Global Burden of Disease Study, 1990 and 2015. *Rev Bras Epidemiol*, 20, 142–156.
- McKinnon, S.A. Municipality-level estimates of child mortality for Brazil: a new approach using Bayesian Statistics. [Dissertation]. The University of Texas at Austin, 2010.
- McLachlan, G., & Basford, K. (1988). *Mixture models: inference and applications to clustering*. New York: Marcel Dekker.
- McLachlan, G. J., & Peel, D. (1996). An algorithm for unsupervised learning via normal mixture models. *ISIS: Information, Statistics and Induction in Science*, DL Dowe, KB Korb, and JJ Oliver (Eds.), 354–363.
- McLachlan, G. J., Peel, D., Basford, K. E., & Adams, P. (1999). The EMMIX software for the fitting of mixtures of normal and t-components. *J Stat Software*, 4(2), 1–14.
- Mello Jorge, M. H. P., Laurenti, R., & Gotlieb, S. L. D. (2010). Avaliação dos sistemas de informação em saúde no Brasil. *Cad Saúde Colet*, 18, 07–18.
- Moura, E. C., Gomes, R., Falcão, M. T. C., Schwarz, E., Never, A. C. M., & Santos, W. (2015). Gender inequalities in external cause mortality in Brazil, 2010. *Ciênc Saúde Coletiva*, 20(3), 779–788.
- Murray, C. J. L., Rajaratnam, J. K., Marcus, J., Laakso, T., & Lopez, A. D. (2010). What can we conclude from death registration? Improved methods for evaluating completeness. *PLoS Med*, 13, 7(4).
- Murray, J., de Castro Cerqueira, D. R., & Kahn, T. (2013). Crime and violence in Brazil: Systematic review of time trends, prevalence rates and risk factors. *Aggression Violent Behav*, 18(5), 471–483.
- Oliveira, G. L., Loschi, R. H., Assunção, R. M. (2017). A random-censoring Poisson model for underreported data. *Stat Med*, 36(30):4873–4892. <https://doi.org/10.1002/sim.7456>.
- Paes, N. A. (2005). Avaliação da cobertura dos registros de óbitos dos estados brasileiros em 2000. *Rev Saúde Pública*, 39(6), 882–890.
- Paes, N. A. (2007). Qualidade das estatísticas de óbitos por causas desconhecidas dos Estados brasileiros. *Revista de Saúde Pública*, 41(3), 436–445.
- Peralta, A., Benach, J., Borell, C., Espinel-Flores, V., Cash-Gibson, V., Queiroz, B. L., & Mari-Dell'Olmo, M. (2019). Evaluation of the mortality registry in Ecuador (2001–2013)—social and geographical inequalities in completeness and quality. *Popul Health Metrics*, 17(1), 3.
- Pereira, F. N. A., & Queiroz, B. L. (2016). Diferenciais de mortalidade jovem no Brasil—a importância dos fatores socioeconômicos dos domicílios e das condições de vida nos municípios e UFs. *Cadernos Saúde Pública*, 32(9).
- Queiroz, B. L., Freire, F. H. M. D. A., Gonzaga, M. R., & Lima, E. E. C. D. (2017). Completeness of death-count coverage and adult mortality (45q15) for Brazilian states from 1980 to 2010. *Rev Bras Epidemiol*, 20, 21–33.
- Ram, U., Jha, P., Gerland, P., Hum, R. J., Rodriguez, P., Suraweera, W., ... Gupta, R. (2015). Age-specific and sex-specific adult mortality risk in India in 2014: analysis of 0–27 million nationally surveyed deaths and demographic estimates from 597 districts. *Lancet Glob Health*, 3(12), e767–e775.
- Reichenheim, M. E., De Souza, E. R., Moraes, C. L., de Mello Jorge, M. H. P., Da Silva, C. M. F. P., & de Souza Minayo, M. C. (2011). Violence and injuries in Brazil: the effect, progress made, and challenges ahead. *Lancet*, 377(9781), 1962–1975.
- Rodrigues, N. C. P., Monteiro, D. L. M., Almeida, A. S. D., Barros, M. B. D. L., Pereira Neto, A., O'Dwyer, G., ... Lino, V. T. S. (2016). Temporal and spatial evolution of maternal and neonatal mortality rates in Brazil, 1997–2012. *J Pediatra*, 92(6), 567–573.
- Ruther, M., Leyk, S., & Buttenfield, B. (2017). Deriving Small Area Mortality Estimates Using a Probabilistic Reweighting Method. *Ann Am Assoc Geographers*, 1–16.
- Schmertmann, C. P., & Gonzaga, M. R. (2018). Bayesian estimation of age-specific mortality and life expectancy for small areas with defective vital records. *Demography*, 55(4), 1363–1388.

- Silva, B. G. C. D., Lima, N. P., Silva, S. G. D., Antúnez, S. F., Seerig, L. M., Restrepo-Méndez, M. C., & Wehrmeister, F. C. (2016). Mortalidade materna no Brasil no período de 2001 a 2012: tendência temporal e diferenças regionais. *Rev Bras Epidemiol*, 19, 484–493.
- Soares Filho, A. M., Cortez-Escalante, J. J., & França, E. (2016). Revisão dos métodos de correção de óbitos e dimensões de qualidade da causa básica por acidentes e violências no Brasil. *Ciência Saúde Colet*, 21, 3803–3818.
- Souza, A., Hill, K., & Dal Poz, M. (2010). Sub-national assesment of inequality trends in neonatal and child mortality in Brazil. *Int J Equity Health*, 9, 21.
- Souza, M. F. M., França, E. B. & Cavalcante, A. (2017). Carga da doença e análise da situação de saúde: resultados da rede de trabalho do Global Burden of Disease (GBD) Brasil. *Revista Brasileira de Epidemiologia*, 20(supl. 1), 1–3.
- Stephens, A. S., Purdie, S., Yang, B., & Moore, H. (2013). Life expectancy estimation in small administrative areas with non-uniform population sizes: application to Australian New South Wales local government areas. *BMJ Open*, 3(12), e003710.
- Szwarcwald, C. L., de Frias, P. G., Júnior, P. R. B. D., de Almeida, W. D. S., & de Moraes Neto, O. L. (2014). Correction of vital statistics based on a proactive search of deaths and live births: evidence from a study of the North and Northeast regions of Brazil. *Population health metrics*, 12(1):16.
- Tabutin, D., Masquelier, B., Grieve, M., & Reeve, P. (2017). Mortality Inequalities and Trends in Low-and Middle-Income Countries, 1990–2015. *Population*, 72(2), 221–296.
- Tsimbos, C., Kalogirou, S., & Verropoulou, G. (2014). Estimating spatial differentials in life expectancy in Greece at local authority level. *Popul Space Place*, 20(7), 646–663.
- Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. *Appl Latent Class Anal*, 11, 89–106.
- Vermunt, J. K., & Magidson, J. (2000). *Latent GOLD user's guide*. Boston: Statistical Innovations. Boston: Innovations Inc.
- Waiselfiz, J. J. (2013). *Mapa da violência 2013: homicídios e juventude no Brasil*.
- Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., et al. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the global burden of disease study 2015. *Lancet*, 388(10053), 1459–1544.
- Ward, M. D., and Gleditsch, K. S. (2007). An introduction to spatial regression models in the social sciences. *Manuscript at* <http://www.faculty.washington.edu.mdw>. Last visited August, 8, 2007.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)